

運動機能に支援を必要とする人をサポートする スマートフォン制御システムの試作

宮崎英一, 坂井 聡*, 谷口公彦**, 野田知良***,
大野香織****, 篠原智代****
(技術教育)(特別支援教育)*, (香川県立高松養護学校教諭)**,
(香川県立聾学校教諭)***, (かがわ総合リハビリテーションセンター)****

760-8522 高松市幸町1-1 香川大学教育学部
*760-8522 高松市幸町1-1 香川大学教育学部
**761-8057 高松市田村町1098 香川県立高松養護学校
***761-8074 高松市太田上町513-1 香川県立聾学校
****761-8057 高松市田村町1114番地 かがわ総合リハビリテーションセンター

The trial production of a smart phone control system that Motor dysfunction.

Eiichi MIYAZAKI, Satoshi SAKAI, Kimihiko TANIGUCHI, Tomoyoshi NODA,
Kaori OONO and Tomoyo SHINOHARA

Faculty of Education, Kagawa University, 1-1 Saiwai-cho, Takamatsu 760-8522

** Faculty of Education, Kagawa University, 1-1 Saiwai-cho, Takamatsu 760-8522*

*** Kagawa Prefectural Kagawa Special education school, 1098 Tamura-cho, Takamatsu 761-8057*

**** Kagawa Prefectural School for the Deaf, 513-1 Otakami-machi, Takamatsu 761-8074*

***** Kagawa General Rehabilitation Center, 1114 Tamura-cho, Takamatsu 769-2302*

要旨 本研究では、運動機能に支援を必要とする人に対してスマートフォンの制御を音声入力でサポートするシステムを試作した。このシステムでは、ゲーム機として一般的に市販されているキネクトを音声認識エンジンとして使用する事で、安価で正確に音声認識が行える事が出来た。更に音声認識においては1つの制御命令に対して複数の該当単語を割り当てる事がで

きるので，事前学習が不必要となり直感的に使用できるシステムとなった。また，音声発信の制御以外にも，音量の調節等が通話中に音声入力によってリアルタイムで制御できるので，利用者の使用環境の改善が見込まれ，結果的に利用者の情報生活の質的向上が期待できる事も示された。

キーワード スマートフォン，音声認識，キネクト，adb，制御

1. はじめに

現在，通話機能を有する家庭用の小型通信端末は従来の携帯電話（いわゆる，フィーチャーフォンあるいはガラパゴス・ケータイとも称されるタイプを指す。本論文では，これらの通信機器を以後フィーチャーフォンと称する。）からスマートフォンに置き換えられている。これは総務省の保有率調査¹（主な情報通信機器の普及状況（世帯））を見ると平成22年度：9.7%，平成23年29.3%，平成24年49.5%と毎年大幅に増加しており，今後スマートフォンが一般的な民生利用においても情報通信機器の主流となるのは間違いないであろう。

これらの通信機器を運動機能に障害をもった方が利用する事を想定した場合，携帯電話とフィーチャーフォンでは，ユーザインタフェースにおいて，大きな差異が認められる。従来のフィーチャーフォンでは小さいボタンではあるが，機械的な接点を持つ電話番号入力キーを備えており，これが物理的なボタン・キーとして実装されていた。このため，押した感触が操作者に指を通じてフィードバックされるので，指先の感触で押されたことが理解できるという大きな特徴を持っていた。よってフィーチャーフォンは，運動機能に障害のある方でも，比較的使いやすい入力システムとなっていた。

さらにこれらは，内部ハードウェアの実装としてキー入力が直接的に電話番号に割り当てられていたため，フィーチャーフォンの使用状況によらず，キーが押されると，それが直ちに電話番号として入力されるという利点もあった。

一方，スマートフォンにおいてはタッチパネルを入力インタフェースとしたものが大部分であり，これが直感的操作方法をユーザに提供している。このため，初心者にも操作方法の習得が簡単であり，誰でもが利用しやすいインタフェースという利点を生かして広く一般に普及してきた。しかし，タッチパネルは物理的に平面状のスイッチであり，それを押した時に凹むという従来のスイッチに相当する感覚のフィードバックが無く，運動機能に障害のある方にとっては押したかどうかという感覚を掴みにくい入力インタフェースであるといえる。さらに視覚障がいを持つ方にとっては，ボタンの位置を触覚で認識できないので，ここでも入力インタフェースとして問題になっていた。

そこで本研究ではスマートフォンとコンピュータをUSBで接続し，キネクトの音声認識システムを用いて，通話等のスマートフォンの制御を行うシステムを試作した。本研究で試作したシステムでは，音声通話だけに留まらず，音声だけで音量の変更や特定のプロ

グラムの起動、インターネットブラウザの制御等を行う等の様々な制御を行う事が可能となる。

よって本システムを利用する事で、運動機能に障害のある方でも、従来では困難であったスマートフォンの、より幅広い利用が可能になる。この結果、利用される方に対して日常生活における情報活用の質的向上が十分期待できるものと考えられる。

2. 制御システムの概略

現在、スマートフォンとして幾つかの機種が発売されているが、日本で販売されている代表的なものとしては1) Android端末、2) iOS端末、3) Windows Phone端末があげられる。このなかで、本研究では1)のAndroid端末を制御対象とした。これは国内において多くの台数が販売されており、数多くの実生活環境内で利用可能な事、コンピュータからのハードウェアの制御が比較的に簡単に行える等の理由によるものである。以後、本研究ではスマートフォンと称するものはAndroid端末を指すものと定義する。

本研究では図1に本システムの概略を示す。ここでは音声認識システムとしてマイクロソフト社から販売されているキネクトを入力インタフェースとして応用している。キネクトの想定される最も一般的な使用方法

は、キネクトのカメラで操作対象の人物を撮影し、その動きを入力インタフェースとしてNUI (Natural User Interface) を提供するものである。

本来、このシステムはゲーム用の入力インタフェースとして開発されたが、そのNUIの持つ優れた直感的な入力特性を生かして東京大学の中村研究室でオークシステム^{2, 3}, (Observation and Access with Kinect) 等、障がい者の支援に応用した例⁴も報告されている。

本研究では、NUIの核となるカメラ部分を利用せず、このシステム内に実装された音声認識機能を入力インタフェースとして利用し、自作プログラムからスマートフォンの制御を行うものである。ここで従来から研究されてきた音声認識システムを本システムに適用する場合、以下の3種類の音声認識エンジンの利用法が考えられる。

- 1) ローカルマシン－音声認識エンジン
- 2) ネットワーク－音声認識エンジン
- 3) キネクト－音声認識エンジン

ここで1)の場合にはシステム全体の構成が、コンピュータ本体+マイクで構成されるために、全体としてコンパクトなシステムとなる。しかし音声認識の処理にコンピュータの高い情報処理能力が要求されるため、ある程度のCPUパワーを持ったコンピュータの利用が要求される。よってこのシステムでは、



Kinect (音声認識) ⇔ [USB] ⇔ コンピュータ ⇔ [USB] ⇔ スマートフォン

図1 制御システム概略図

比較的高価なコンピュータを準備する必要がある。

2) のシステムでは1) と同様にコンピュータ本体+マイクだけで構成されるため，これもコンパクトなシステムとなる。しかしネットワーク上の音声認識エンジンを利用するため，必ずネットワーク環境が必要となり，ネットワークを介さないオフラインでの運用が困難である。この問題は多くの家庭においては問題にならない。しかし，恒常的に病院等の施設に入院している利用者を想定した場合には，個人で院内でのネットワークの構築は難しい場合がある。またネットワークを利用している事から，ネットワークの混雑状況によっては認識結果がタイムラグを持って表示される場合も有る。

3) で提案した本システムの場合，コンピュータ本体以外にキネクトが必要なため，1)，2) と比較するとシステム全体は大きなシステムとなる。しかし1)，2) の運用においてはマイクが必須になるため，システム的环境によっては音声を適切な入力レベルまで増幅するためのアンプ等の付加装置が必要になる場合もある。

一方，キネクト本体には入力用のマイクアンプも予め準備されているので，別途これらを準備する必要が無いという利点もある。このように考えるとキネクトを利用する本システムは1つのデバイスの準備だけで完成するので，運用上において，使いやすいシステムであると言える。

3. 音声認識

3.1 音声コマンド

音声認識でスマートフォンの制御を行う場合，予め音声とスマートフォンの制御命令を関連させておく必要がある。この制御命令を

表1 登録された音声コマンド

音声	コマンド
みやざき、ぎじゅつ	087-832-***4 電話発信
さかい、とくし	087-832-***5 電話発信
たにぐち、ようご	087-832-***6 電話発信
あげる	音量上げる
さげる	音量下げる
きる、おわる	通話終わる

音声コマンドとする。本システムで実装した音声コマンドを表1に示す。ここでは，表1に示した複数の音声認識の単語を1つの音声コマンドとして認識させているので，従来の音声認識システムと比較して，より柔軟な入力が可能である。

この音声コマンドの割り当ては状況に応じて任意に増加させる事が可能なので，利用者の目的に応じて自然な入力を提供する事が可能となる。このため利用者は事前に音声コマンドの学習を行う事なく，簡単に本システムの運用が開始できる。よって病院内のように事前学習が困難な環境下の利用者にも負担をかけないシステムとなっている。

3.2 音声認識の流れ

本システムのフローを図2に示す。具体例として，キネクトに向かって音声「宮崎」を発声したとする。この音声キネクトによって「みやざき」という単語データとして抽出され，プログラム内で定義された音声コマンドと連携される。その音声コマンドを利用して，制御プログラムとは別の外部プログラム(cmd.bat)が子プロセスとして実行⁵される。ここでは音声コマンドが宮崎のスマートフォ

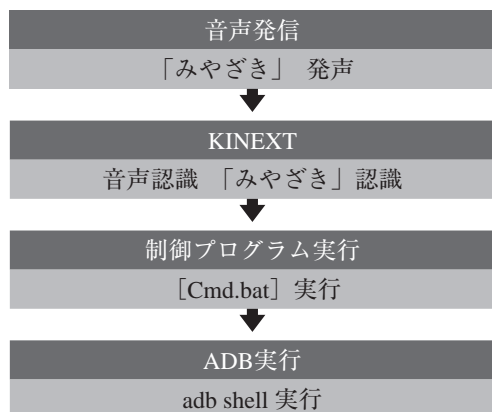


図2 プログラムのフロー

ンに電話を発信させており、この命令をファイル内 (cmd.bat) に記述している。具体的には

```
{ adb shell am start -a
android.intent.action.CALL tel:電話番号}
```

と1行記載しているだけである。

ここでこの (cmd.bat) 内に記載された電話番号をメモ帳等のエディターソフトを用いて書き換えてやれば、任意の電話番号に対して音声を利用した発信が可能になる。このため病院や家庭内等の実際の運用では、プログラムの開発環境の準備が不要になり、だれでもが簡単に自分の目的に合わせて病院や学校といったその場でのプログラムのカスタマイズが可能になるという利点がある。

3.3 スマートフォンの制御

キネクトを用いて音声認識を行う場合、本研究ではコンピュータからスマートフォンの制御を行うのにadb⁶ (Android Debug Bridge) を利用している。本来、adbはコンピュータからAndroid端末 (スマートフォン) のデバッグを行うのに使用されており、Android

のSDK (Software Development Kit) に標準 (platform-tools) で含まれている。よって、コンピュータにこの開発環境をインストールする事で、誰でもがadbの利用が可能になる。例えば、コンピュータに接続されているデバイス (スマートフォン) を確認する場合は、「adb devices」というコマンドを発行すると

```
# adb devices
# List of devices attached
# 06580843f0de0a01      device
```

のような結果が得られる。ここで表示された「06580843f0de0a01」がコンピュータに接続された実機 (スマートフォン) のIDを示している。また、これ以外にも「adb logcat」とすれば

```
# adb logcat
----- beginning of /dev/log/system
D/MobileDataStateTracker( 743):
default: Received state=CONNECTED, old=CONNECTED,
reason=(unspecified)
D/MobileDataStateTracker( 743):
default: Broadcast received: android.intent.
```

とコンピュータに接続されたデバイスの動作ログを確認する事もできる。本研究ではこのデバック機能を利用してコンピュータからスマートフォンの通話や制御を行っている。

本研究においては上記で説明したように、音声通話の制御は

```
#adb shell am start -a
android.intent.action.CALL
tel:xxxxxxxxxxxxx
```

とインテントで電話番号を発信制御している。(画面の構成上、3行で記載しているが、本来は1行の連続したプログラムである)。

本画面のxxxは通話先の電話番号を示しており，この番号を変更する事で，任意の相手と音声通話を行う事が可能になる。さらに，このコマンドを実行した場合，スマートフォンの画面は直接電話発信開始の状態になるので，1 ボタンの操作のみで通話が可能になる。本研究では，このコマンドを通話先に応じて，予め幾つか作成しておき，これを音声認識プログラムから呼び出す事で複数の相手に応じ音声通話が可能になる。これ以外にも

```
#adb shell input keyevent  
KEYCODE_VOLUME_DOWN
```

のコマンドを実行すればスマートフォンの音量大小の制御も可能になる。よって，周囲の環境に応じてスマートフォンの制御も行えるので，実際の日常生活において使用しやすい環境の構築が可能になる。

図3に本研究で試作した音声認識システムを用いてスマートフォンから音声発信を行っている画面を示している。同図a)では，キネクトにより認識された音声コマンドを示している。ここでは発声した「みやざき」という単語が認識され，プログラムから外部通話プログラムが実行されている。

この音声コマンドが実行されれば，同図b)に示したスマートフォンの制御が行われる。ここでは上記で説明したように，確認画面を経ずに，直接通話状態まで画面が遷移している事が示されている。このため，複雑な手続きを経ず，通話が可能になるので，ボタンを2度押しさせないような利用者に負担をかけるシステムとなっている。

これ以外にも通話中にキネクトに向かって「あげる」と発声すると，通話をしながらでも音量の調整が可能になり，利用者の利便性の向上が可能になった。



図3 b) 電話番号発信中



図3 a) [みやざき] 発声を認識

4. おわりに

本研究では，ゲーム機器の入力インターフェースとして開発された音声認識システムを用いて，コンピュータからスマートフォンの音声発声を制御するシステムを試作した。ここではスマートフォンの制御にはAndroid SDKに含まれるadbを用いて通話制御を行っている。このため多くの機種で制御が可能になり，多くのスマートフォンで実行が可能になると考えられる。

今後は，本システムを実際の環境で使用して頂き，日常生活で使われる音声コマンド等の追加・修正等を行い，日常生活の質的向上を目指すものである。

5. 謝辞

本研究は，平成26年度科学研究費補助金（基盤研究（C））「運動機能及び発達障害をサポートする生活・学習支援ワンデバイスシステムに関する研究」（課題番号24500648）の一

部として行われたことを記して謝意を示す。

6. 参考文献

¹ 総務省トップ＞政策＞白書＞25年版＞主な情報通信機器の普及状況（世帯）

<http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h25/html/nc243110.html>

² 企業市民活動（企業の社会的責任）ホーム＞事例紹介＞「OAK（オーク）」障害のある方の活動支援を1歩進めたソリューション

<http://www.microsoft.com/ja-jp/citizenship/story/201211.aspx>

³ 重度障害児向けプログラム - TOP＞OAKの開発

<http://doit-japan.org/doat/oak/>

⁴ Kinect 白杖システムによる上り階段の認識と基礎的評価、山口翔太郎、滝沢穂高、青柳まゆみ、江崎修央、水野慎士、情報処理学会インタラクション 2013、pp.746-749, 2013

⁵ キネクトを用いた障がい者支援用入力デバイスの試作、宮崎英一、坂井聡、谷口公彦、野田知良、大野香織、篠原智代、香川大学研究紀要報告第Ⅱ部、pp.*-*, 2014

⁶ Android Debug Bridge

<http://developer.android.com/tools/help/adb.html>